

(19)日本国特許庁 (JP)

(12)公開特許公報 (A)

(11)特許出願公開番号

特開平11-225310

(43)公開日 平成11年(1999)8月17日

(51)Int.C1.

識別記号

H04N 5/93
5/76
5/91

F I

H04N 5/93
5/76
5/91

Z
B
N

(21)出願番号

特願平10-25824

(22)出願日

平成10年(1998)2月6日

審査請求 未請求 請求項の数 5 O L (全8頁)

(71)出願人 000004226

日本電信電話株式会社
東京都新宿区西新宿三丁目19番2号

(72)発明者 谷口 行信

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72)発明者 阿久津 明人

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72)発明者 外村 佳伸

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(74)代理人 弁理士 鈴木 誠

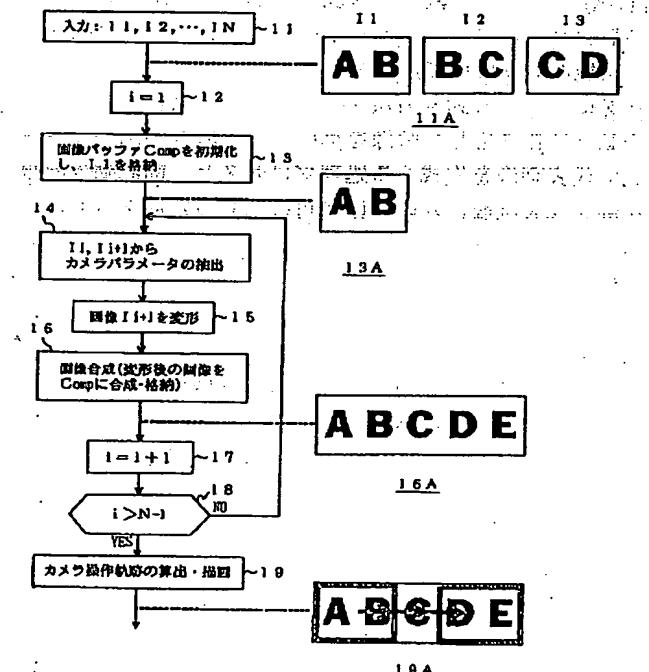
最終頁に続く

(54)【発明の名称】映像静止画表示方法及び装置並びに映像静止画表示プログラム格納記録媒体

(57)【要約】

【課題】 映像データベースシステム、映像編集システムにおいて、連続した映像を静止画表示する際、映像の動き情報を利用者が直感的に理解できる形で表示する。

【解決手段】 連続した映像1.1Aを入力する工程(1 1)、該画像に含まれる動き情報を抽出する工程(1 4)、該動き情報に応じて、映像を構成する複数枚の画像を合成し、一枚の代表画像1.6Aを生成する工程(1 5, 1 6)、動き情報を代表画像と対応づけて可視化し描画1.9Aする工程(1 9)を設ける。



【特許請求の範囲】

【請求項1】 連続した映像を静止画表示する方法であつて、映像に含まれる動き情報を抽出する第1の工程と、前記動き情報に応じて、映像を構成する複数枚の画像を合成し、一枚の代表画像を生成する第2の工程と、前記動き情報を前記代表画像と対応づけて可視化する第3の工程を具備することを特徴とする映像静止画表示方法。

【請求項2】 請求項1記載の映像静止画表示方法において、第1の工程が動き情報としてカメラまたは被写体の動作速度を抽出し、第3の工程が前記動作速度を可視化することを特徴とする映像静止画表示方法。

【請求項3】 請求項1または2記載の映像静止画表示方法において、第3の工程が映像に含まれる動きを分類して、分類結果に応じて予め設定された複数の可視化方法のうちの一つを選択することを特徴とする映像静止画表示方法。

【請求項4】 連続した映像を入力する映像入力手段と、前記映像に含まれる動き情報を抽出する動き情報抽出手段と、前記動き情報に応じて、映像を構成する複数枚の画像を合成し、一枚の代表画像を生成する画像合成手段と、前記動き情報を前記代表画像と対応づけて可視化する動き情報可視化手段と、前記画像合成手段と前記動き情報可視化手段から得られる画像や情報を出力する映像出力手段と、を具備することを特徴とする映像静止画表示装置。

【請求項5】 連続した映像を静止画表示するための映像静止画表示プログラムを格納したコンピュータ読み取り可能な記録媒体であつて、映像に含まれる動き情報を抽出する処理プロセスと、前記動き情報に応じて映像を構成する複数枚の画像を合成して代表画像を作成する処理プロセスと、前記動き情報を前記代表画像と対応づけて可視化する処理プロセスを含むことを特徴とする映像静止画表示プログラム格納記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は映像データベースシステム、デジタル映像編集システムなどのユーザインターフェースを構築する際に、連続的に撮影された映像を静止画表示する方法および装置並びにそのプログラムを格納した記録媒体に関する。

【0002】

【従来の技術】 ビデオデッキを使ってビデオテープの特定の場面を探し出すには、通常、早回しや巻き戻しを繰り返すしか手法が無く、時間と手間がかかるという問題があった。

【0003】 これを解決するための一つの手法として、映像の場面の変り目(カット)を検出し、場面毎に1枚の静止画を場面を代表する画像(代表画像と呼ぶ)として抽出し、それを一覧表示する方法が知られている。この方法によれば、映像を時間順に見なくても、代表画像を一覧するだけで所望の場面を視覚的に短時間で検索できる効果がある。しかし、この方法には次のような問題点がある。

【0004】 (1) パン(カメラを水平方向に回転させる操作)、チルト(カメラを垂直方向に回転させる操作)などのカメラ操作を含むシーンでは、シーンの全体像を表現する代表画像を抽出することが困難である。例としてカメラを下から上に動かしながら人の足元から顔までを撮影した映像を考える。この場面の最初の画像には足元しか写っていないので、シーンの一部しか写っていないという意味で、代表画像としては不適切である。し、最後の画像には顔しか写っていないので、これもまた代表画像としては不適切である。

(2) 映像から抽出された代表画像にはカメラ操作のタイプ、方向、速度が表現されていない。なお、カメラ操作のタイプとしては、パン、チルトのほかに、ズーム(ズームレンズの画角を変化させて被写体サイズを連続的に変える操作)、トラック(カメラを左右方向に移動させる操作)、クレーン(カメラを上下方向に移動させる操作)などがある。

(3) 映像から抽出された代表画像だけでは被写体の動きが表現できない。

【0005】 (1) の問題点を解決する一つの方法として、連続する複数枚の画像を合成して広視野の画像を生成し、それを代表画像として用いるアプローチがある(例えば、特開平5-304675号公報)。これは、映像の中でカメラがどのように動いたかを画像処理によって算出し、そのずれの量だけ画像を平行移動、拡大縮小しながら合成することによって、1枚の広視野画像を生成するというものである。この方法は(1)の問題点を解決するが、(2)や(3)の問題点は解決されない。

【0006】 (2) を解決する一つの手法として、カメラ操作のタイプと方向を矢印表現し、映像から抜き出された代表画像に重畳して表示する方法がある。(「認識技術を応用した対話型映像編集方式の提案」(電子情報通信学会論文誌、D-I I, Vol. J 75-D-1 1, No. 2, pp. 216-225))。図2に、この方法による表現の一例を示す。図2の21、22、23はそれぞれパン、チルト、ズームを矢印表現している。

【0007】 しかし、この従来技術ではカメラ操作の速度を表現できていないし、(1)や(3)の問題点を解決していない。さらに、矢印が画像内容と対応づけられないため、カメラ操作によって画像内容がどう

に変化するかが分からぬことが問題である。即ち、カメラ操作前の画像は代表画像として表示されるが、カメラ操作後の画像が表示されないからである。また、カメラ以外の動き情報（被写体の動き情報）の表現方法については述べられていない。

【0008】

【発明が解決しようとする課題】本発明は、上記従来技術の問題点を解決するためになされたものであり、カメラの動きがあつてもシーン全体を表現する代表画像を生成し、カメラ操作のタイプ、方向、速度などを代表画像に対応づけて表示し、さらに、被写体の動きを直感的に理解できるように代表画像に対応づけて表示する、映像静止画表示方法及び装置、並びに、そのプログラム格納記録媒体を提供することを目的とする。

【0009】

【課題を解決するための手段】請求項1の発明の映像静止画表示方法は、映像に含まれる動き情報を抽出し、該動き情報に応じて、映像を構成する複数枚の画像を合成し、一枚の代表画像を生成し、動き情報を代表画像と対応づけて可視化することを特徴とする。映像に含まれるカメラ操作に代表される動き情報を抽出し、その情報に基づいて画像を合成することで、カメラの動きがあつてもシーン全体を表現する代表画像を作成できる。

【0010】請求項2の発明は、請求項1の映像静止画表示方法において、動き情報としてカメラまたは被写体の動作速度を抽出し、該動作速度を可視化することを特徴とする。

【0011】請求項3の発明は、請求項1または2の映像静止画表示方法において、映像に含まれる動きを分類して、分類結果に応じて予め設定された複数の可視化方法のうち一つを選択することを特徴とする。

【0012】請求項4の発明の映像静止画表示装置は、連続した映像を入力する映像入力手段と、該映像に含まれる動き情報を抽出する動き情報抽出手段と、該動き情報に応じて、映像を構成する複数枚の画像を合成し、一枚の代表画像を生成する画像合成手段と、動き情報を代表画像と対応づけて可視化する動き情報可視化手段と、画像合成手段と動き情報可視化手段から得られる画像や情報を出力する画像出力手段とを具備することを特徴とする。

【0013】請求項5の発明の映像静止画表示プログラム格納記録媒体は、連続した映像を入力し、そのシーンを代表する静止画を表示するアルゴリズムにおいて、映像に含まれる動き情報を抽出する処理プロセスと、前記動き情報に応じて映像を構成する複数枚の画像を合成して代表画像を作成する処理プロセスと、前記動き情報を前記代表画像と対応づけて可視化する処理プロセスとを含むことを特徴とする。

$$(x', y') = (ax + b, ay + c)$$

(1) 式のモデルは、ある画像点 (x, y) と直前の画像の点 (x', y') の対応関係を表す。ここで a ,

【0014】

【発明の実施の形態】以下、本発明の一実施形態について図面により説明する。図1は、本発明の映像静止画表示方法の一実施形態の処理フロー図である。図1の処理フローの各工程は、画像列を入力するための映像入力装置、ハードディスクやそれに準ずる装置、動き情報を算出するため、あるいは画像合成するために必要となる画像バッファのメモリーやそれに準ずる装置、静止画を表示するためのディスプレイなどの表示装置を備え、それらハードディスク、バッファメモリ、表示装置を予め定められた手順に基づいて制御するコンピュータ(CPU)やそれに準ずる装置により、適宜、実行することができる。この種のコンピュータシステムは周知であるので、図示は省略する。

【0015】図1において、ステップ11は連続的に撮影されたN枚の画像 I_1, I_2, \dots, I_N を入力する。入力映像が複数の場面から構成されている場合には、予め場面の切り替わりを検出し、ショットの単位に分割しておき、ステップ11ではショット毎に画像列を入力する。ステップ12は画像の番号を示す変数 i を1に初期化する。ステップ13は画像バッファ(Comp)を初期化し、最初の画像 I_1 を画像バッファに格納する。ステップ14は隣り合う2枚の画像 I_i, I_{i+1} から動き情報をとしてカメラパラメータを抽出する。カメラパラメータはカメラの動きを定量的に表すものであり、その抽出方法の詳細については後述する。ステップ15はカメラパラメータに従って画像 I_{i+1} を変形し、ステップ16は変形後の該画像を画像バッファCompに既に格納されている画像と合成して格納する。11Aに示した画像列の例では、カメラを左から右へパンしているので、ステップ15では画像 I_{i+1} を逆に左に平行移動する変形を行い、ステップ16では、該変形後の画像を順次それまでの画像と合成することにより16Aに示す横長の合成画像を得る。ステップ17は変数 i に1を足し、ステップ18は画像列をすべて処理し終わったか検査し、終わらない場合、ステップ14に戻り、終ったならステップ19へ進む。ステップ19は、ステップ14で得られた各カメラパラメータに基づいて、カメラの動きを示す軌跡を後述する方法で算出し、画像バッファの画像に軌跡を上書き描画して、最終的な静止画像19Aを表示する。

【0016】ここで、ステップ14のカメラパラメータの抽出方法について説明する。カメラパラメータを抽出するためには、まず、カメラモデルを定義する。カメラモデルは、2次元画像中の点が次の時間の画像中のどの点に移動するかを示す数学的モデルであり、本実施形態ではカメラのパン、チルト、ズームを考慮した次の(1)式のモデルを用いる。

(1)

50 像の点 (x', y') の対応関係を表す。ここで a ,

b, cをカメラパラメータと呼び、それぞれカメラのズーム、パン、チルトに対応している。例えば、カメラ操作がズームせず右方向にパンしている場合、 $a=1$, $b>0$, $c=0$ となる。カメラがズームインしている場合、 $a<1$ となる。カメラモデルとして(1)式で定義されるもの以外にも、平行移動モデル、アフィンモデル、Bilinearモデル、射影モデルなどを用いることができる。

$$\begin{aligned} \text{MSE}_i(a, b, c) &= \sum_{x, y} \{I_{i+1}(x, y) - I_i(x', y')\}^2 / N \\ &= \sum_{x, y} \{I_{i+1}(x, y) - I_i(ax + b, ay + c)\}^2 / N \end{aligned} \quad (2)$$

【0019】ただし、和($\sum x, y$)は、すべての対応する画素について計算され、Nは対応する画素の数である。誤差量の評価関数としては、MSE以外にもロバスト統計の考え方を導入したM推定法などもある。

【0020】誤差量MSEを最小とするa, b, cが求

$$\text{MSE}_i(a_i, b_i, c_i) = \min_{a, b, c} \text{MSE}_i(a, b, c) \quad (3)$$

【0022】(3)式で定義される最適化問題の解法としては、単純にa, b, cを変化させながらMSE_i(a_i, b_i, c_i)を算出し、その最小値を達成するパラメータを出力する単純な方法でもよいが、処理コストが膨大であることが問題である。

【0023】この解決策としては、解像度を段階的に落とした画像(ピラミッド画像と呼ばれる)を作成し、まず解像度の最も低い画像を使ってカメラパラメータを推定し、推定されたカメラパラメータを初期値として、もう一段階解像度の高い画像について最適解を得る、ということを段階的に解像度を高めながら繰り返すことによってパラメータを効率的かつ精確に推定する方法がある(H. S. Sawhney and S. Ayer: Compact Representation of Videos Through Dominant and Multiple Motion Estimation, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vo

$$\begin{aligned} W_k(x, y) &= (x', y') \\ &= (a_i(a_i(\dots(a_i(x+b_{i-1})+b_{i-2})+\dots \\ &\quad b_1, a_i(a_i(\dots(a_i(y+c_{i-1})+c_{i-2})+\dots+c_1)) \end{aligned} \quad (4)$$

と表すことができる。ただし、便宜的にW₀(x, y) = (x, y)と定義する。

$$T_{r a j} = \{W_0(W/2, H/2), W_1(W/2, H/2), \dots, W_{N-1}(W/2, H/2)\} \quad (5)$$

となる。ただし、W, Hはそれぞれ画像の幅と高さを表す。

【0028】カメラ操作軌跡は、点列T_{r a j}を折れ線

【0017】二枚の画像I_i, I_{i+1}から、その間のカメラの動きを表すカメラパラメータa_i, b_i, c_iを以下のようにして推定する。二枚の画像I_i, I_{i+1}の誤差量を見積るために平均二乗誤差MSE (Mean Square Error) を次の(2)式のように定義する。

【0018】

【数1】

【0019】ただし、和($\sum x, y$)は、すべての対応する画素について計算され、Nは対応する画素の数である。誤差量の評価関数としては、MSE以外にもロバスト統計の考え方を導入したM推定法などもある。

【0020】誤差量MSEを最小とするa, b, cが求めるべきカメラパラメータa_i, b_i, c_iであると考える。すなわち、

【0021】

【数2】

1. 18, No. 8, pp. 814-830, 1996)。

【0024】ステップ15の画像変形操作は、(1)式に上記のようにして求めたカメラパラメータa_i, b_i, c_iを代入した式を用いて行う。

【0025】以上説明した方法は、画像処理によってカメラパラメータを推定するものであったが、カメラに取り付けられたセンサによって動きが測定されてカメラパラメータが取得可能である場合にはそれを用いてよい。

【0026】次に、ステップ19のカメラ操作軌跡の算出・描画手続きについて述べる。式(1)より、k=1, 2, ..., N-1について、画像I_{k+1}上の点(x, y)と画像I₁上の点(x', y')の対応関係は

【0027】(4)式から、画像バッファComp内の座標系で、画像中心点の軌跡を表す点列T_{r a j}は

(5)

と矢印を用いて描画することによって表わすことができる。パン、チルトの場合には、それぞれ図3の31, 32に示すような図を得られる。なお、カメラ操作の開始

点、終了点を示すために、19Aに示したように、図1で開始画像と終了画像を囲む矩形を描画することも好適である。

【0029】カメラ操作軌跡の描画は、図3の37に示すように、折れ線の終点にのみ矢印を描画するようにしてもよいが、この場合にはカメラ操作の速度が表現されないことが問題である。カメラ操作の速度を表現する目的には、例えば図3の31に示すように、矢印を一定時間間隔で描画し、矢印の長さが速度に対応するような表現を行うことが好適である。

【0030】次に、映像に含まれる動きを分類し、その結果に応じて動き情報の可視化方法を選択する手続きに

$$Z_i = a_1 \times a_2 \times \cdots \times a_{i-1}$$

として算出する。ステップ52は、カメラ操作がズームを含むか否かを判定する。具体的には、ズーム率 Z_i が、すべての $i = 1, 2, \dots, N$ について $Z_{min} < Z_i < Z_{max}$ を満たす場合にはズーム操作を含まないと判定する。ただし、 Z_{min}, Z_{max} は予め与えられた閾値であり、 $Z_{min} < 1$ かつ $Z_{max} > 1$ を満たすように設定される。カメラ操作がズームを含まない場合には、上で説明したように、ステップ53で画像の中心点の軌跡を算出・描画する。カメラ操作としてズームを含む場合には、ステップ54で、画像の四隅の軌跡を同様に(4)式を用いて算出しその軌跡を描画する。

【0032】ステップ54により、ズームイン、ズームアウトの映像についてそれぞれ図3の33, 34に示すような静止画像が得られる。また、ズームしながらチルトしているシーンでは、図3の35に示すような静止画像が得られる。図3の36は、カメラが光軸の周りに回転している場合の図示方法の一例である。

【0033】なお、上記実施形態ではカメラ操作軌跡を矢印を用いて描画したが、その他にも直線の明暗によって方向を示したり、矢印の太さで速度を表したりすることもできる。また、ズーム操作の有無によって、表現方法を切り替える例について示したが、例えば、カメラがパンしている映像において、被写体をフォローしているシーンか否かを判定して、被写体の動きの可視化方法を切り替えるようにすることもできる。

【0034】図4は、本発明の映像静止画表示方法の他の実施形態の処理フロー図である。これも、実際にはコンピュータの支援のもとに実現される。図1ではカメラの動きを動き情報として可視化するものであったが、図4では、被写体の動きを動き情報として抽出する。

【0035】図4において、ステップ41は連続する画像列を入力する。本実施形態では、固定されたカメラによって撮影された映像 I_1, \dots, I_N (41A)を入力として仮定する。もちろん、図1と同様にしてカメラの動きを考慮することも可能である。

【0036】ステップ42は、最初の画像 I_1 について被写体領域 Obj_1 を指定する。ステップ43は、 $k =$

について述べる。カメラがパン、チルトしている場合は、図3の31, 32のように画像中心点の軌跡でよいが、ズームの場合には画像中心点は変化しないのでズームの向き(ズームインかズームアウトか)を可視化できない。この問題点を解決するためには、「カメラ操作がズームを含むか否か」でカメラの動きを分類・判別し、判別結果に応じて表現方法を切り替えることが有効である。これを図5の処理フロー図を用いて説明する。

【0031】図5において、ステップ51は、カメラ操作を分類するために必要となるズーム率 Z_i ($i = 1, 2, \dots, N$)を

(6)

$2, 3, \dots, N$ について画像 I_k を順次調べて被写体の動きを追跡する。具体的には、被写体領域 Obj_{k-1} と最も相関の高い領域を画像 I_k 内から検索し、それを画像 I_k の被写体領域 Obj_k として出力するということを、 $k = 2, \dots, N$ について繰り返す。図4の43Aはこれを示す。相関の算出方法としては、周知の色ヒストグラム差分、輝度相関など様々なものを用いることができる。

【0037】ステップ44は、最初と最後の被写体領域 Obj_1, Obj_N が浮き出るように画像 I_1, I_N を合成することによって、図4の44Aに示すような合成画像を得る。具体的には、合成画像 $Comp(x, y)$ を次式に従って合成する。

$$(i) \text{ 点 } (x, y) \text{ が } Obj_1 \text{ にも } Obj_N \text{ にも含まれない} \\ Comp(x, y) = \{ I_1(x, y) + I_N(x, y) \} / 2$$

$$(ii) \text{ 点 } (x, y) \text{ が } Obj_1 \text{ に含まれる} \\ Comp(x, y) = I_1(x, y) \\ (iii) \text{ 点 } (x, y) \text{ が } Obj_N \text{ に含まれる} \\ Comp(x, y) = I_N(x, y).$$

【0038】ステップ45は被写体動き軌跡を算出して描画する。被写体動き軌跡は、被写体領域 Obj_i ($i = 1, \dots, N$)の重心の軌跡として求めることができる。この被写体動き軌跡を合成画像 $Comp(x, y)$ 上に描画して、最終的な画像46Aをディスプレイ等に表示する。

【0039】図6は、本発明の映像静止画表示装置の一実施形態のブロック図である。図6において、映像入力手段61は、映像信号を入力し、映像の切れ目を検出して、連続したシーン毎に画像列を動き情報抽出手段62に渡す。動き情報抽出手段63は、複数枚の画像をバッファに保持し、図1や図4の実施形態で説明したような方法を用いてカメラの動きあるいは被写体の動きを抽出する。画像合成手段64は、抽出された動き情報に基づいて画像列を合成し、内部バッファに保持する。動き情報可視化手段64は、動き情報を、同じく図1や図4の実施形態で説明したような方法を用いて可視化する。画

像出力手段6.5は、映像毎に画像合成手段6.4と動き情報可視化手段6.5から得られる画像を重畠してディスプレイあるいはコンピュータスクリーン内のウィンドウ、あるいはプリンターに出力表示する。

【0040】図6で説明した映像静止画表示装置の各部の手段は、図1や図4の処理フローの各工程と同様に、画像列を入力するための映像入力装置、ハードディスクやそれに準ずる装置、動き情報を算出するため、あるいは画像合成を行うために必要となる画像バッファとしてメモリーやそれに準ずる装置、静止画を表示するためのディスプレイなどの表示装置とを備え、それらハードディスク、バッファメモリ、表示装置を予め定められた手順に基づいて制御するコンピュータ(CPU)やそれに準ずる装置により、適宜、実行することが可能である。

【0041】以上、本発明の実施形態について説明したが、図1や図4の処理フロー、図6の各部の機能は、映像静止画表示プログラムとしてコンピュータが読み取り可能な記録媒体、例えばフロッピーディスクやメモリーカード、MO、コンパクトディスク(CD-ROM)、テープなどに記録して提供することが可能である。

【0042】また、以上の実施形態では、一つの映像を静止画表示する場合について述べたが、複数シーンから成る映像を映像に分割し、それについて実施形態で説明した方法や装置によって静止画を作成し、複数枚の該静止画を一覧表示するようにすることも、映像内容を短時間で把握可能とする目的のために好適である。

【0043】

【発明の効果】本発明によれば、動き情報に応じて画像を合成して代表画像を作成し、該代表画像と対応づけて

動き情報を可視化するようにしたので、カメラの動きがあっても映像の全体を表現する代表画像を抽出でき、また、カメラ操作のタイプ、方向、速度を代表画像に対応づけて表示することができ、さらには、被写体の動きを代表画像に対応づけて表示できる効果がある。その結果として、映像にまつわる動き情報を利用者が直感的に理解できるような形で表示することができるようになるので、映像内容を効率的に把握でき、目的の映像を視覚的に短時間で検索できるようになる。

【図面の簡単な説明】

【図1】本発明の一実施形態例を示す処理フロー図である。

【図2】従来技術におけるカメラ操作可視化方法を説明するための模式図である。

【図3】本発明による静止画表示例を示すための模式図である。

【図4】本発明の他の実施形態例を示す処理フロー図である。

【図5】動き情報可視化手続きの一実施形態例を示す処理フロー図である。

【図6】本発明の更に他の一実施形態例の構成を示すブロック図である。

【符号の説明】

6.1 映像シーン入力手段

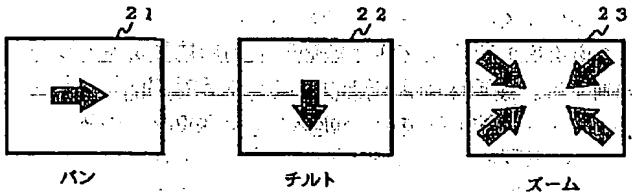
6.2 動き情報抽出手段

6.3 画像合成手段

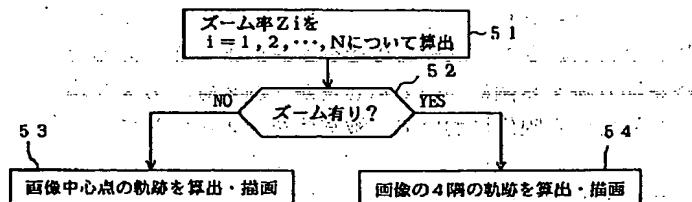
6.4 動き情報可視化手段

6.5 画像出力手段

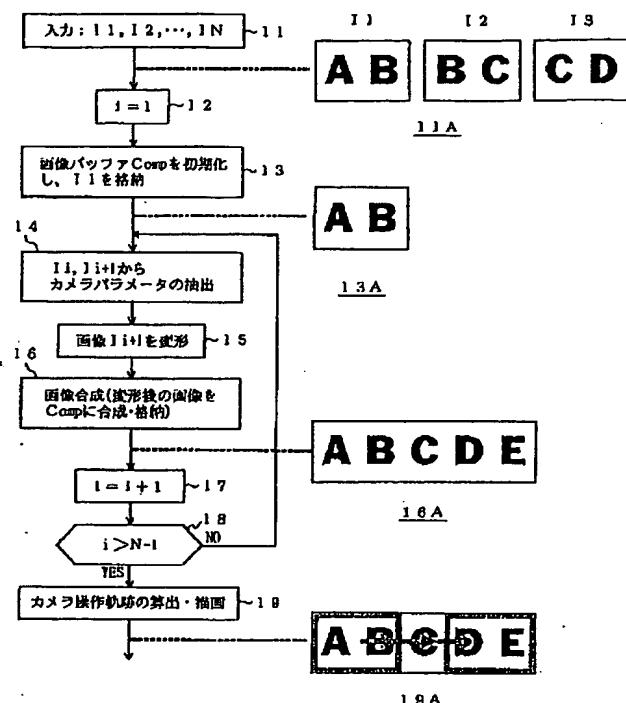
【図2】



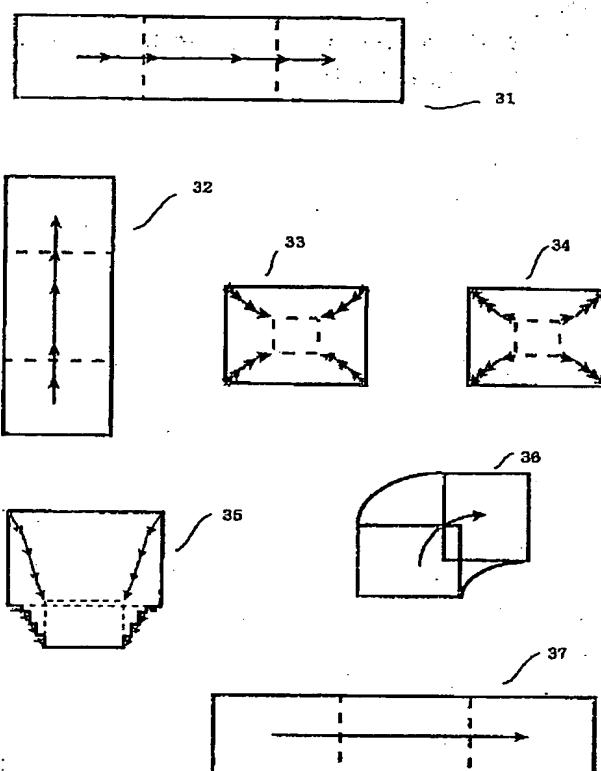
【図5】



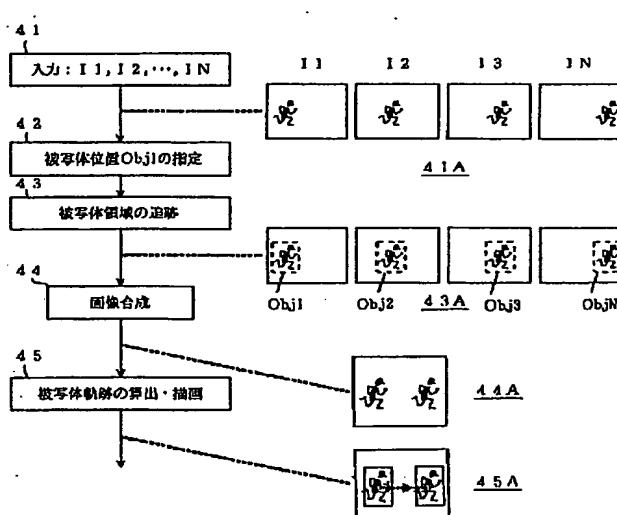
【図1】



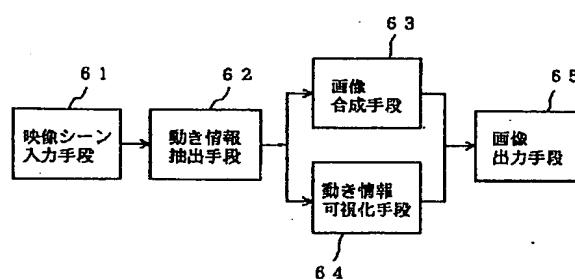
【図3】



【図4】



【図6】



フロントページの続き

(72)発明者 佐藤 隆

東京都新宿区西新宿三丁目19番2号 日本

電信電話株式会社内